图书情报工作
第 66 卷 第 13 期　2022 年 7 月

国际科研数据政策导向解析

■ 秦顺[1]　戴柏清[2]
[1] 武汉大学信息管理学院　武汉 430072　[2] 中国人民大学信息资源管理学院　北京 100872

摘　要：[目的/意义]剖析国际科研数据政策的导向及其特征，为我国完善科研数据政策保障体系、优化科研数据管理与共享工作提供重要依据。[方法/过程]采用案例分析、文献调研方法归纳出国际科研数据政策的“整合 DLC-SH”和“机器可操作”两种导向，并使用模本法分析、验证两种导向的异同。[结果/结论]“整合 DLC-SH”导向秉持科研数据管理与共享整体观，遵循“组织管理 - 业务流程”基本框架，目的在于构建全球协同式数据治理方略；“机器可操作”导向提供依次递进的指导性原则，重视数据基础设施的辅助作用，目的在于实现“数据互联网”全球布局。两种导向在理论基础、核心组件和服务对象上存在不同，但在基础环境、实现途径和最终目标上基本一致，且两种导向的相互融合成为现阶段的主流趋势。建议我国强化政策的导向功用，融入国际工作体系；重塑工作机制与流程，完善数据治理格局；“人文与技术”相结合，构建“人机协同”生态。

关键词：科研数据政策　数据生命周期　利益相关者　机器可操作　FAIR 原则
分类号：G301
DOI：10.13266/j.issn.0252 - 3116.2022.13.005

1　引言

数据密集型科研范式下，科研数据管理与共享成为国际范围内高度关注的问题。国际科研数据管理与共享工作呈现出两种主要动向，且与科研数据政策导向密切关联。一方面，数据具有自然和社会属性，自然属性指向数据生命周期(Data Life Cycle，DLC)维度，社会属性则指向利益相关者(Stakeholder，SH)维度，通过整合数据生命周期与利益相关者两个维度来推动科研数据管理与共享工作发展成为主要趋势，本文将其统称为“整合 DLC-SH”导向。例如，《澳大利亚高校/机构科研数据管理政策大纲》[1]、英国《五步制定科研数据政策》[2]和瑞士《科研数据管理政策模板》[3]等代表性的政策大纲/模板均是从“整合 DLC-SH”的角度构建的；我国国务院办公厅于 2018 年颁行的《科学数据管理办法》[4](以下简称《办法》)也建立起了“国家统筹、各部门与各地区分工负责”的制度设计，形成了以“数据生命周期为经、利益相关者为纬”的工作机制[5]。另一方面，克服数据发现与重用障碍迫切需要推动科研数据管理与共享的机器可操作，本文将其称为“机器可操作”导向。FAIR 原则便是其中影响较为广泛的政策指南，其强调提高机器自动查找和使用数据的能力以及支持它在重复使用方面的作用[6]。

在“整合 DLC-SH”和“机器可操作”两种科研数据政策导向下，国际组织、政府机构、研究与资助机构、图书馆等利益相关者的科研数据管理与共享行动得到有序推进，积极践行这两种政策导向的欧美等发达国家和地区的科研数据管理工作体系也日趋完善。相比之下，我国的科研数据政策保障体系较为滞后，亟待依据国情并参考国际科研数据政策导向，优化科研数据管理与共享工作的制度设计。

2　科研数据政策研究概况

2.1　研究现状

国内外科研数据政策研究已取得较为丰硕的成果[7]。对科研数据政策的探究主要面向国际组织、政府及其各部门、科研资助者、科研机构、高校、出版机构、数据中心等不同利益相关者；覆盖计划、管理、共享和利用等不同生命周期流程；关涉海洋、人口与健康、生命科学、医学、农业、天文等不同学科领域。从研究主题来看，既有科研数据政策的解读与优化研究、政策框架及评价指标体系研究，更不乏政策内容的调研分

作者简介：秦顺，博士研究生，E-mail：qinshun_xtu@163.com；戴柏清，博士研究生。
收稿日期：2021 - 11 - 21　修回日期：2022 - 01 - 14　本文起止页码：48 - 60　本文责任编辑：易飞

析与借鉴。

除此之外，亦有一些机构和学者从国际视角对科研数据政策的实践经验进行总结。国际数据委员会(CODATA)与其成员合作，通过调查国际科研数据政策最佳实践，确定关键要素，为全球政策制定与实施提供基础框架[8]；F. C. Correa da Silva[9]分析了科研数据管理的基础设施服务和发展政策，提出国际通用的科研数据管理与共享机制。我国学者更多地关注对科研数据政策框架、要素和内容的剖析，研究集中在以下方面：①对国际科研数据政策及指南等文本的解读，如对《科研数据管理国际联盟实用指南》[10-11]、科学数据管理与共享的 FAIR 原则[12-14]、联机计算机图书馆中心(OCLC)《科研数据管理的现实》系列报告[15]等国际政策、准则与指南的介绍，为我国科研数据管理与共享工作提供参考。②对国际科研数据政策的比较分析，如尤霞光等[16]以经济合作与发展组织、联合国教科文组织、国际科学联盟、欧盟委员会等国际组织为例，进行对比分析；盛小平等[17]通过分析国际组织科学数据开放共享政策，归纳出不同利益相关者在科学数据开放共享中的责任与作用。③对国际科研数据政策法规体系的解构，如唐义等[18]通过解析国际组织及一国政府、科研资助机构、期刊社及科研机构等主体的科学数据共享政策，认为国际科学数据共享的政策法规体系由宏观、中观、微观三个层次构成，其奠定了关联科学的制度基础。

2.2 研究述评

分析现有研究可以发现，“整合 DLC-SH”和“机器可操作”两种导向已深刻影响到了国际科研数据政策的研究与实践。但是，相关研究仅停留在内容解析、对比和经验析出层面，缺乏对科研数据政策全球实践的系统梳理和规律总结。基于此，本文尝试归纳出国际科研数据政策的两种导向，并分析其特点和异同，以期为我国进行政策设计和推动实践发展提供依据。

3 研究设计

3.1 研究样本来源

一方面，为保证研究样本的代表性，笔者设定了两个选取标准：①样本须为国际重要科研联盟发布的宣言、指南和原则等指导性政策文本，由于现有研究和实践中仍无统一的国际科研数据政策清单可供借鉴，在进行样本选择时笔者参考了尤霞光等[16]、盛小平等[17]和邢文明等[19]列出的重要国际组织，并据此收集其最新政策文本；②为析出国际科研数据政策的最新导向，将政策颁行的时间窗口定为 2015 年 1 月 -2021 年 7 月。如表 1 所示，重点收集了 2015 年以来发布的较新且具有典型性的国际科研数据政策，为本文的研究样本。

表 1 国际科研数据政策样本基本情况

政策代码		政策名称	政策制定者	颁布日期	政策导向
SE		科研数据管理国际联盟实用指南[20]	科学欧洲协会	2021 年 1 月	IDS/MA
RDA		数据存储库的 TRUST 原则[21]	研究数据联盟	2020 年 5 月	IDS/MA
CODATA		科研数据北京宣言[22]	国际科学理事会数据委员会	2019 年 11 月	IDS/MA
GIDA		土著数据治理的 CARE 原则[23]	全球土著数据联盟	2019 年 9 月	IDS
ERC		科学出版物与科研数据开放获取实施指南[24]	欧洲研究理事会	2017 年 4 月	IDS/MA
OCLC		科研数据管理的现实[25]	联机计算机图书馆中心	2017 年 3 月	IDS
EC	EC_1	2020 计划框架下的科学出版物与科研数据开放获取规则指南[26]	欧盟委员会	2017 年 3 月	IDS/MA
	EC_2	2020 计划框架下的 FAIR 数据管理指南[27]		2016 年 7 月	IDS/MA
SI		大数据世界中的开放数据[28]	国际科学联盟	2015 年 12 月	IDS/MA
ICSU-WDS		WDS 数据共享原则[29]	国际科学理事会世界数据系统	2015 年 11 月	IDS

注：①IDS-Integrated DLC-SH 的缩写，即“整合 DLC-SH”导向；MA-Machine Actionable 的缩写，即“机器可操作”导向；②表中政策按颁布日期降序排列，下同

另一方面，为科学解析国际科研数据政策的导向，需经过理论与实践互证的验证过程。为此，在广泛收集国际科研数据政策的基础上，跟进最新国际科研数据实践发展动态，重点收集了世界经合组织(OECD)、国际科学联盟(SI)、国际科学理事会(ICSU)、欧盟委员会(EC)、欧洲研究理事会(ERC)、科学欧洲协会(SE)、CODATA、研究数据联盟(RDA)、世界数据系统(WDS)、国际图联(IFLA)和 OCLC 等组织的政策实践案例，以及 FAIR 原则、TRUST 原则等国际通用准则的实践，为验证国际科研数据政策导向提供双重证据。

3.2 研究方法与过程

通过对所采集的国际科研数据政策及实践案例的内容分析，将国际科研数据政策导向归纳为“整合 DLC-SH”和“机器可操作”两种，并运用比较研究法分析其主要特征。在此基础上，使用叶鹰[30]提出的模本法对比两种国际科研数据政策制定与实施导向机制的异同。模本法是采用两类具有不同特征的极小样本(样本数≤6)，经主体和客体两方面分析达成一致结论的双证法，其操作过程是：①选取主体模本，提取特征；②选取客体模本，提取特征；③主客体特征指向一致时结论有效。模本法提供了“理论与实践互证”的典型途径，能够有效适配本文所提供的国际科研数据政策文本(理论)、实践案例(实践)双重证据，其中，样本政策文本为主体模本，实践案例为客体模本。

4 “整合 DLC-SH”的国际科研数据政策导向

4.1 秉持科研数据管理与共享整体观

数据生命周期是指从产生数据开始，经过对数据的收集、加工、分析、存储、访问和共享，最终实现数据再利用的循环过程[31]；“利益相关者”一词最早见于斯坦福研究院的一份内部备忘录中，系指“支持某组织生存必不可少的群体”[32]。科研数据管理与共享的整体观主张使数据生命周期各个环节、系统内部利益相关者相互关联，机构与外部环境也相互影响，构成一个整体生态系统[15]。有学者指出，科研数据管理服务处于由人员、政策、资源和技术组成的系统中，在研究人员和机构产生、收集、使用和保存数据时，可为他们提供支持和指导[33]。

如表2所示，通过分析国际科研数据政策样本库中规范的科研数据管理与共享生态系统构成，可以发现：在“整合 DLC-SH”导向下，数据生命周期、利益相关者和基础环境是科研数据管理与共享工作有机整体的关键组成部分，其中基础环境包括基础设施(如软件、工具和平台等)、法律政策、资金支持和数据管理计划(Data Management Plan，DMP)模板等支撑性要素。

分析表明，国际科研数据政策基本秉持科研数据管理与共享的整体观。例如，SE、CODATA、ERC、OCLC、EC_1 和 EC_2 等6项政策规范的业务流程完整地覆盖了数据生命周期；SE、RDA、CODATA、OCLC、EC_2、SI 和 ICSU-WDS 等政策或设置专章来明晰各利益相关者的权责分工，或分别为不同的利益相关者提供实践指南；大多数的政策文本皆明确提出要为推动科研数据管理与共享工作的有序发展配置一定的基础环境。其中，政策 OCLC[25]充分体现了数据管理的整体观，其建立了科研数据管理 E-E-C 模型并引入由数据素养教育、专业知识服务、数据监护支持构成的数据服务簇概念，E-E-C 模型分别指向用户、馆员和技术三个维度，其中专业知识服务和数据素养教育是“软件”，而数据监护支持则是“硬件”，形成了一个整体生态系统。

4.2 遵循“组织管理－业务流程”基本框架

“整合 DLC-SH”的国际科研数据政策导向指向的利益相关者和数据生命周期两条主线，实质上反映了科研数据管理与共享工作的组织管理和业务流程两个维度[7]。科研数据管理与共享工作具有复杂性、动态性和持续性等特点[31]，必须构建“横纵联动，协同合作”“分级管理，分工负责”的管理体制，打通从计划、管理、共享到利用的业务流程。

4.2.1 理顺管理机制，明确责任主体及其职责

如表3所示，以本文的样本政策为证据源，通过归纳其对组织管理维度的规范，发现与盛小平等[17]的研究相近，国际科研数据政策中大多明晰了科研数据管理与共享工作的责任主体及其职责，并构建了有序的工作机制。这些政策厘定的责任主体包括政府机构、科研组织、研究与资助机构、专业协会/学会与学术团体和图书馆或档案馆等，它们分别在科研数据管理与共享工作中扮演着领导者、引导者、管理者、组织者、资助者和协调者等角色，并形成由“制度设计、凝聚共识、资源分配、连接枢纽、协调推广、管理服务、保存共享、工具支持、控制评估、生产管理、共享利用”构成的有机管理体制和责任分工。各类责任主体分别从宏观、中观和微观层面出发，为完善政策法规、优化业务流程和供给多元服务等提供有效支持，力图构建起职责明晰的科研数据管理与共享工作体系。

4.2.2 覆盖数据生命周期，优化核心业务流程

如表4所示，依据前期研究成果[34]，笔者设计了由计划、管理、共享和利用组成的4个维度、8个观测点，对样本政策规范的核心业务流程进行归纳分析，数据表明：共计6项样本国际科研数据政策覆盖了全数据生命周期，仅3项政策未提及 DMP 编制和实施，政策 RDA[21]的作用主要是为了提高可信存储库评估和选择的效能。可以发现，实现有序、良好的数据生命周期管理及其流程优化思想已融入到国际科研数据政策制定与实施过程中，其目的在于联结各利益相关者共同促使科研数据管理与共享工作形成一个不断生长着的有机体，打造数据生态链，促进信息生态的演进发展和良性循环。

表2　样本政策中科研数据管理与共享整体观的表现

政策代码		生态系统构成		
		DLC	SH	基础环境
SE		DMP 核心要求、撰写指南和评估准则第 1-3、5 条和可信存储库的最低标准 1-4 条等皆依据数据生命周期组织	①DMP 核心要求、可信存储库的最低标准适用于研究与资助机构,DMP 撰写指南、可信存储库选择指南适用于科研人员,DMP 评估准则适用于评审员; ②DMP 核心要求、撰写指南和评估准则第 6 条为“数据管理职责与资源”	DMP 核心要求、撰写指南和评估准则第 4 条提出“法律和道德要求、行为准则”的标准;DMP 核心要求、撰写指南和评估准则第 6 条规范了确保满足 FAIR 原则所需的资源
RDA		主要面向数据创建、存储阶段,主张提供可信存储服务	单独列出“责任”一条主要原则,面向的主要利益相关者为数据中心、软件与工具制造商、科研机构、科研人员等	提供可信存储的基础设施、技术支持
CODATA		十条原则之一是“DMP 及政策制定是必要的”,且其内涵涉及数据访问、发现和重用等流程	阐明了推进相关领域多边合作的核心原则,核心利益相关者是国际组织、政府机构等	实施需要清晰的政策和指导方针、专项经费支持以及明确的承诺
GIDA		在整个数据生命周期中将人和目的放在首位,应用于数据收集、集成、分析	CARE 原则(CARE Principle,即集体利益-Collective Benefit、控制权-Authority to Control、责任-Responsibility 和道德-Ethics),以解决土著人民对其数据的权利和利益为目的	①强调在全数据生命周期阶段和整个数据生态系统中,土著人民的权利和福祉应是首要关注的问题; ②支持发展数据劳动力和数字基础设施,以便能够创建、收集、管理、利用数据,保障安全、优化数据治理
ERC		规定参加开放科研数据先导计划(ORD 试点)需要三个步骤,并采取积极措施促使第三方能够访问、挖掘、利用、复制和传播科研数据	创建“选择退出”机制,参与者主要为地平线 2020 资助的科研机构与科研人员	通过 ORD 试点、DMP 模板、科研数据管理工具与技术平台等实现科研数据的开放获取
OCLC		创新性运用了数据服务簇概念,其由数据素养教育、专业知识服务、数据监护支持等科研数据服务组成,贯穿全数据生命周期	主要聚焦于总结研究型大学科研数据管理的理论、内容、因素与策略,为全球提供参考	①使用指向用户、馆员和技术的科研数据管理服务的三维分析框架(E-E-C)界定了科研数据管理服务的边界; ②考察了支持新兴数据管理实践所需的基础设施、服务和其他资源
EC	EC_1	主要强调科研数据的开放获取,即以“开放为常态,不开放为例外”(As Open as Possible, as Closed as Necessary)	同 ERC 的规范	通过 ORD 试点,保障科研数据开放获取的权利
	EC_2	认为“DMP 是科研数据管理中的关键要素,贯穿于整个科研过程和数据生命周期”,并提供包括数据概述、数据的 FAIR 原则、资源配置、数据安全、数据伦理及其他问题等核心在内的 DMP 模板	利益相关者如研究人员、专业出版商、软件与工具制造商、资助机构、公共科学机构等	①所采纳的 FAIR 原则是为解决现有数字生态系统的阻碍而设计的一套简明扼要的、可测量的、增强其科研数据可重用性的原则; ②实现 FAIR 原则的支撑条件包括资金支持、DMP 模板等
SI		倡导的实践为:数据引用、互操作、非限制性重用、可关联性,重点放在管理与利用阶段	在开放数据的原则中宣示其责任者包括科学家、科研机构与高校、出版商、资助机构、专业协会、学术团体与组织、图书馆、档案馆和存储库等	评估了数据“开放的边界”,例外情况限于隐私、安全和公共利益的要求
ICSU-WDS		①政策的目的声明为:促进普遍、公平地获得质量可靠的科学数据、数据服务、产品和信息,以期实现长期数据管理; ②强调充分开放共享数据、元数据、产品和信息	利益相关者为国际组织、科学社区(学会、协会)、数据生产者、数据用户和数据管理者等	遵守国家、国际法律和政策以及国际研究道德行为标准

表3　样本政策对组织管理维度的规范

责任主体	角色定位	核心职责	证据源
政府机构	领导者、资助者、生产者、出版者、传播者、管理者和利用者	制度设计:科研数据政策制定、管理体制设计、资助数据开放共享、推动全数据生命周期的科研数据管理与共享等	CODATA、ICSU-WDS
科研联盟	引导者	凝聚共识:凝聚全球/区域/行业科研数据管理与共享的共识,提出指导性原则,发布实施科研数据政策、规范、标准和指南等	SE、CODATA、GIDA、EC_2、ICSU-WDS
科研资助机构	资助者	资源分配:承担分配公共/资助资金等职责,具备影响其支持的科研机构的整体数据政策和管理制度、科研人员的数据管理与共享意愿的作用	SE、ERC、OCLC、EC_1、EC_2、SI
科研教育机构	生产者、管理者、传播者和利用者	连接枢纽:构建良好的科研数据管理与共享环境、协同上级与横向主体完善并实施科研数据政策,会同数据中心、图书馆等开发信息通讯技术与基础设施,协调科研资助机构保障数据安全、辅助科研数据与 DMP 质量评估,开展数据服务	ERC、EC_1、SI、ICSU-WDS

（续表3）

责任主体	角色定位	核心职责	证据源
专业协会/学会、学术团体	推广者和协调者	协调推广：制定科研数据政策和准则，协调、推广科研数据开放共享活动	SI、ICSU-WDS
图书馆或档案馆、数据中心	组织者、管理者和服务提供者	管理服务：提供覆盖全数据生命周期、多元数据服务簇的科研数据管理与共享服务，进行科研数据管理与共享技术工具、平台的研发应用等	SI、ICSU-WDS
数据出版商	出版者	保存共享：实现科研数据的可信存储，出版与共享（元）数据，提供验证科学主张的数据等	EC_2、SI
软件与工具制造商	开发者和运营者	工具支持：主导信息通讯技术与基础设施研发、运营，如DMP工具、数据平台、存储库等技术平台和工具资源的开发和应用	EC_2
评审员	评估者	控制评估：对科研数据、DMP以及数据服务的科学和技术质量进行评估	SE
科研人员	生产者、传播者、管理者和利用者	生产管理：遵守科研数据政策、法律和伦理道德要求，编制DMP并积极融入数据生命周期管理流程等	SE、ERC、EC_1、EC_2、SI、ICSU-WDS
数据用户	传播者和利用者	共享利用：实现科研数据的免费共享和高效利用	GIDA、ICSU-WDS

表4 样本政策规范的核心业务流程

政策代码		业务流程
SE		P→U
RDA		M
CODATA		P→U
GIDA		M→U
ERC		P→U
OCLC		P→U
EC	EC_1	P→U
	EC_2	P→U
SI		M→U
ICSU-WDS		M→U

注："P"表示计划阶段，主要关涉DMP制度设计和实施；"M"表示管理阶段，核心任务是数据创建、存储；"S"表示共享阶段，重点业务是数据共享、出版；"U"表示利用阶段，面向科研用户开展数据分析、重用、评估与验证服务；"→"表示核心业务从起点到终点的全流程指向

4.3 构建全球协同式数据治理方略

管理属于微观"运营"范畴，治理则属于宏观"制度"范畴，两者存在一定程度上的重合与关联。数据治理更强调去中心及多元主体协同参与，目标是通过利益均衡和科技向善实现治理平衡，打造各主体相互协同、良性互动的治理生态。数据共享的核心要义在于协同，全球科研数据管理与共享的协同共治已取得良好成效。如表5所示，政策解析和实践皆表明，串联起"数据生命周期和利益相关者"的全球协同式数据治理方式成为科研数据管理与共享工作的发展之道。例如，政策SE[20]业已被法国国家研究机构（ANR）、瑞典研究委员会（VR）和荷兰研究委员会（NWO）等8个机构所共用，在欧洲形成广泛影响，对我国的启示是要制定科研数据管理政策并推动落实、充分发挥科研联盟机构的合作优势以及理顺数据管理与FAIR原则的关系等[11]。

5 "机器可操作"的国际科研数据政策导向

5.1 提供依次递进的指导性原则

"机器可操作"是国际科研数据政策的另一大导向。早在2014年，国际性学术社区FORCE11[38]发布的重要性、信誉与归属、论据、唯一标识、便于访问、长久性、明确性与可验证性、互操作性与灵活性等8条数据引用共同原则即体现了这一导向。近年来，国际上陆续提出的FAIR原则[6]（FAIR Principle，即可发现－Findable、可访问－Accessible、可互操作－Interoperable和可重用－Reusable）以及TRUST原则[21]（TRUST Principle，即透明度－Transparency、责任明晰－Responsibility、以用户为中心－User focus、可持续发展－Sustainability和技术性支持－Technology）等均是一系列依次递进的指导性原则。FAIR原则、TRUST原则的前一条子原则皆是后一条的前置条件，各子原则环环相扣，为机器可操作提供了更多可能性；TRUST原则为FAIR原则的实现奠定了基础，两者互相关联，如TRUST原则中的R子原则指向FAIR原则的A1子原则，S子原则直接指向F、A、R子原则。现阶段，FAIR原则与TRUST原则在全球范围内的应用已较为成熟。

5.1.1 科研数据管理与共享的FAIR原则[6]

FAIR原则以克服数据发现与重用障碍为追求[39]，包含了4个递进关系的科学数据处理阶段，体现了科学数据融合从低层到高层的融合深度[40]；实现步骤包括检索不符合FAIR原则的数据、对检索到的数据进行分析、定义语义模型、使数据可链接、设计许可协议、为数据集定义元数据和数据资源发布等[41]，形成了层层递进的机器可操作规范。目前，FAIR原则已

表5 国际科研数据政策的适用范围与实施进展

政策代码		适用范围	实施进展列举
SE		全球	①SE 的 8 个成员机构已经使用该政策来制定或调整其机构中的 DMP 要求,其他机构则专注于 DMP 要求的开发与实施,该指南已成为校准 DMP 和科研数据政策的依据; ②下一步,将采取措施全面落实各项政策,涉及 DMP 提交、申请缺失的 DMP 与监控合规性、制定评估 DMP 的方法等细节;整合科研人员、科研(资助)机构的需求,研发应用支持科研人员的 DMP 在线平台与工具;考虑到特定学科或机构的特殊性,确保满足更多利益相关者的要求; ③第一版指南已迅速被包括欧盟委员会在内的欧洲众多研究基金和执行机构采用,成功地确保了在国家一级的吸收;第二版(DMP 评估准则扩展版),吸引了越来越多的研究基金和致力于改进数据管理的执行机构加入,真正支持整个欧洲和其他地区科研数据管理政策的进一步调整[10]
RDA		全球	已被 4TU. ResearchData、DataONE 和 Dryad 等 39 个技术平台和组织采用[21]
CODATA		全球	秉持鼓励全球合作的核心原则,已被国际部分著名科研机构采纳,如美国国立卫生研究院(NIH)发布的《开放数据政策草案》即符合该宣言的主要原则[35]
GIDA		全球	适用于具有保护数据主权需求的国家和实体。一系列数据集正在努力适配 FAIR 和 CARE 原则,如地球科学信息合作伙伴(ESIP)即制定了相关指南和运营服务的其他标准[36]
ERC		欧盟	吸纳了科研数据管理与共享的 FAIR 原则,强化了 DMP 编制和实施
OCLC		全球	归纳了英国爱丁堡大学、美国伊利诺伊大学厄巴纳-香槟分校、澳大利亚莫纳什大学、荷兰瓦格宁根大学的实践经验,可供全球高校及其图书馆、科研数据管理机构以及科研数据服务部门借鉴[15]
EC	EC_1	欧盟	适用于欧盟研究与创新框架计划"地平线 2020"(Horizon 2020 Programme)下的各利益相关者,框架下的欧洲开放获取基础设施研究项目(OpenAIRE)以及 ORD 试点皆采用该政策,2017 年 1 月开始,ORD 试点已扩展到地平线 2020 计划中所有的专题领域[26]
	EC_2	欧盟	同 ERC,且与 EC_1 相近
SI		全球	SI 的目标之一是改进全球科学政策制定情形,在国际科学政策领域用同一种权威的声音说话[37]。主要倡导数据引用、互操作、非限制性重用、可关联性 4 项实践,被国际社会科学理事会、国际科学院合作组织、国际科学理事会和世界科学院等所采纳[17]
ICSU-WDS		全球	主要被 ICSU-WDS 的普通成员、网络成员、合作伙伴成员和准成员组织采用

被 SE、CODATA 和 EC_2 等样本政策所参考[12];已在欧洲 GO FAIR Initiative 项目、欧盟委员会联合研究中心的 JRC-MRD 多学科科学数据基础设施平台、欧盟委员会 EOSC 计划、美国国立卫生研究院(NIH)"数据共享试验项目"等项目和平台上得到应用。2019 年发布的政策 CODATA[22] 明确指出:"科研数据具有全球公共产品的基本属性,应按照 FAIR 原则的精神推动其开放和共享利用",实施 FAIR 行动在全球得到了积极响应。

5.1.2 可信存储库选择的 TRUST 原则[21]

目前,世界范围内已有超过 2 000 个数据储存库,其中绝大多数未通过认证机构的成熟度和可信度评估[20]。2019 年 4 月,在美国宾夕法尼亚州费城举行的 RDA 第 13 次全体会议上,RDA 成员讨论并提出了 TRUST 原则。TRUST 原则具备辅助 FAIR 原则实现的作用,旨在警醒全球数据存储库使用的利益相关者需要融入基础设施的开发和维护,以促进对数据的持续管理和使用[20]。WDS 及数据认可印章组织(Data Seal of Approval, DSA)参与了 TRUST 原则的编制,其联合推出的认证体系"CoreTrustSeal 认证"即参考、使用了该原则;此外,该原则还得到了 4TU. ResearchData、DataONE 和 Dryad 等 39 个技术平台和组织的采纳[21]。

5.2 重视数据基础设施的辅助作用

FAIR 原则和 TRUST 原则皆较为重视数据基础设施的辅助作用。例如,样本政策 EC_2 较好地适配了 FAIR 原则,向全球科研用户推荐了 Re3Data 登记的数据存储库以及一些允许科研人员自主存放出版物和数据的开放存取数据存储库(如 Zenodo、OpenAIRE 和 CERN 等),同时提供 DMP 工具、链接工具和科学发现平台[27]。CoreTrustSeal 是一个用于评估和验证全球可信数据存储库的认证体系,其构建了"CoreTrustSeal + FAIR"合作机制,认证指标体系包括组织与基础设施、数据对象管理和技术 3 个维度的 16 项指标[42],诸多指标适配了 FAIR 原则(见表 6),并体现了 TRUST 原则的精神,截至 2021 年 8 月,已评估认定 WDS 认证存储库 42 个、DSA 认证存储库 16 个和 CTS 认证存储库 111 个[43],区域布局与全球数据存储库集中在欧美地区的特征基本一致。FAIRsFAIR 是一个旨在促进欧洲的 FAIR 数据实践的组织,目标是为全生命周期的科研数据管理适配 FAIR 原则提供实用解决方案,其支持 FAIR 原则的实践建议如下[44]:①研发和实施数据共享和互操作框架;②为实现 FAIR 原则提供适当的 DMP 支持,并为机器可操作 DMP 制定路线图、指南和工作流程;③为 FAIR 原则的实现提供资金、技术、人员

和基础设施专业支持;④确保数据实现可信监管。在基础设施供给上,FAIRsFAIR 为全球科研数据用户推荐 12 个数据存储库,研发了 FAIR-Aware、F-UJI 等机器可操作的评估指标和工具,以提高科研数据管理与共享工作的互操作性水平。

表 6 CoreTrustSeal + FAIR 认证评估指标[42]

FAIR 原则编号		CoreTrustSeal	
		编号	指标内涵
F	F1	R13	数据发现与标识
	F2		
	F3		
	F4		
A	A1	R15	技术基础设施
	A1.1		
	A1.2	R16	安全
	A2	R10	保存计划
I	I1	R15	技术基础设施
	I2		
	I3	R11	数据质量
R	R1	R11	数据质量
	R1.1	R2	许可
	R1.2	R7	数据完整性与真实性
	R1.3	R15	技术基础设施

5.3 实现"数据互联网"全球布局

"数据互联网"旨在构建一个机器可操作、语义关联、开放共享的全球数据网络,以提高数据的开放性、透明性和可重用性,确保其中的所有数据和相关服务对所有人(包括机器)都是可查找、可访问、可互操作和可重用的。FAIR 原则的使命就是实现"数据互联网"的全球布局,如 WDS 联合 90 多个国际组织共同签署了《赋能 FAIR 数据承诺声明》,共同支持开放和 FAIR 化(FAIRification)数据的目标;B. Mons 发起了 GO FAIR 全球推进计划,目标之一是构建 FAIR 化的数据和服务网络——IFDS(Internet of FAIR Data and Services),其定位为全球性数据基础设施,核心功能是实现数据、工具和计算能力的无缝集成[12]。从机器可操作层面来看,FAIR 原则在实施过程中也进一步体现了构建语义关联的数据互联网的愿景,如表 7 所示,以欧洲为例,FAIRsFAIR 发布的《数据对象评估指标(征求意见稿)》[45]提供了实现科研数据管理与共享工作的机器可操作、语义关联的实施细则,对促使全球科研数据能够被有效发现、获取和使用有积极作用。

表 7 FAIRsFAIR 数据对象评估指标[45]

FAIRsFAIR 编号	评估指标内涵	对应的 FAIR 原则编号
FsF-F1-01D	数据被分配全球唯一的标识符	F1
FsF-F1-02D	数据被分配持久的标识符	F1
FsF-F2-01M	元数据具备描述性的核心元素(创建者、标题、数据标识符、发布者、发布日期、摘要和关键词),以支持数据可发现性	F2
FsF-F3-01M	元数据包括其所描述的数据标识符	F3
FsF-F4-01M	数据以可被机器检索的方式提供	F4
FsF-A1-01M	元数据包含数据的访问级别和访问条件	A1
FsF-A1-02M	元数据可通过标准化通信协议访问	A1
FsF-A1-03D	数据可通过标准化通信协议访问	A1
FsF-A2-01M	即使数据不再可用,元数据仍然可用	A2
FsF-I1-01M	元数据采用合规的知识揭示语言进行描述	I1
FsF-I1-02M	元数据使用了语义资源	I1
FsF-I3-01M	元数据包括数据与其相关实体之间的链接	I3
FsF-R1-01MD	元数据指定了数据的内容	R1
FsF-R1.1-01M	元数据包括数据重用的许可协议信息	R1.1
FsF-R1.2-01M	元数据包括数据创建或生成的来源信息	R1.2
FsF-R1.3-01M	元数据遵循数据目标领域推荐的标准	R1.3
FsF-R1.3-02D	数据以目标领域推荐的文件格式提供	R1.3

通过 FAIR 原则构建"数据互联网"的主要思路为"实现机器可操作基础上的数据标识、关联、集成和可发现,以及智能化排序、检索和利用,从而推动用户与数据的智能连接",关键流程如下:

(1)数据组织。通过数据标识、数据描述来支持数据的可检索和可发现。例如,为数据分配 URL、URN 和数字对象标识符(DOI)等全球唯一且持久的国际化资源标识符(IRI)、统一资源标识符(URI);使元数据

具备描述性的核心元素，譬如依据通用数据引用指南（如 DataCite、ESIP 和 IASSIST 等）以及数据可发现的元数据要求（如 EOSC 数据集最低限度信息、DataCite 元数据架构、W3C 推荐的 Web 最佳实践数据和数据目录词汇表等）确定所需的元数据；规定元数据应包括其所描述的数据标识符、数据以可被机器检索的方式提供，以便用户可以通过元数据发现和访问数据。

（2）数据关联。使用合规的知识揭示语言、语义资源进行元数据描述，并与数据以及其他相关实体之间实现链接，构成可互操作的关联语义网络。例如，使用 RDF、RDFS 和 OWL 等知识揭示语言以及 RDF/XML、RDFa 和 Notation3 等 RDF 序列化格式描述元数据，使用本体论、同义词库、分类学和词汇表等语义资源描述元数据，并通过"相关标识符"（如贡献者 ORCID、出版物 DOI 和机构 ROR 等标识符的协同使用）和"关系类型"等属性将数据与其相关实体联系起来，以便增加其可操作性与重复使用的可能性。

（3）数据发现。限定访问级别与条件，明确可通过标准化通信协议访问（元）数据并保障元数据的永久可用，实现数据的持续可访问。例如，确定公开、不公开、限制或仅限元数据访问 4 类访问级别及其访问条件，建议数据应"以开放为常态，不开放为例外"；给定数据集的标识符，可以使用 HTTP、HTTPS 和 FTP 等标准通信协议检索数据集以及元数据；使用 Datacite 的 PID 服务提供的功能、数据保存计划和策略确保"即使数据不再可用，元数据仍然可用"。

（4）数据利用。提供明确描述数据属性的元数据、推动数据重用的机器可读的许可协议以及有关数据创建或生成来源的信息，并遵循目标领域推荐的标准推动数据的重复使用（如 RDA 提供的元数据原则和标准目录），采用目标领域推荐的文件格式实现数据共享和再利用。

6 国际科研数据政策导向的异同及启示

6.1 国际科研数据政策导向异同

6.1.1 主要差异——基于多特征值的考量

本文采用叶鹰[30]提出的模本法验证"整合 DLC-SH"和"机器可操作"两种国际科研数据政策导向之异同。模本须有主客体一致性和代表性。因此，在差异分析过程中的模本主客体选择标准如下："整合 DLC-SH"情境下，以政策 SE[20] 的 DMP 评估准则扩展版为主体模本，其 DMP 评估准则的实施情况为客体模本；"机器可操作"情境下，以科研数据管理与共享的 FAIR 原则为主体模本，FAIRsFAIR 的实践为客体模本。在此基础上，实现"理论与实践"的互证，具体过程如表 8 所示：

表 8 基于模本法的两种导向差异分析

模本法流程	双证方向	特征提取	
		"整合 DLC-SH"导向	"机器可操作"导向
提取主体模本特征	理论	政策 SE： ①理论基础－数据生命周期与利益相关者理论； ②核心组件－利益相关者（研究与资助机构、科研人员、评审员）、数据生命周期（计划、管理、共享和利用）； ③服务对象－人	FAIR 原则： ①理论基础－关联数据理论、语义网与本体技术； ②核心组件－可发现、可访问、可互操作、可重用； ③服务对象－机器
提取客体模本特征	实践	DMP 评估准则： ①理论基础－数据生命周期与利益相关者理论； ②核心组件－利益相关者（科研人员、评审员）、数据生命周期（数据收集、组织与描述、存储与备份、共享与重用、数据权限与安全等）； ③服务对象－评审员	FAIRsFAIR： ①理论基础－关联数据理论、语义网与本体技术等； ②核心组件－仅未对应 I2 子原则，兼容度达 93.3%； ③服务对象－人机
主客体特征指向	①数据生命周期与利益相关者相结合；②组织管理与业务流程相统一		强调机器可读、可操作
结果	支持：两种导向存在差异		

可以发现，"整合 DLC-SH"和"机器可操作"两种导向在理论基础、核心组件和服务对象等特征上存在一定差别。"整合 DLC-SH"导向更强调人的作用，采纳数据生命周期与利益相关者理论指导科研数据管理与共享工作的有序组织管理与业务流程优化；相比之下，"机器可操作"导向面向机器提供标准规范，强调通过机器可读与可操作推动科研数据的可发现、可访问、可互操作和可重用。

6.1.2 异中求同——两种导向的互相融合

国际科研数据政策的导向趋势是"整合 DLC-SH"和"机器可操作"两者的互相融合。在本文选择的样本政策中，共计有 7 项同时体现了两种导向。如表 9 所示，以政策 SE[20] 规范的 DMP 核心要求（CR）、可信存储库的最低标准（Criterion）指标项与 FAIR 原则的

兼容性为例，两者的指标兼容项为 23 项，兼容率高达 74.19%（23/31）；政策 EC_2[27] 主要通过 DMP 制度设计来规范地平线框架下科研数据管理与共享工作与 FAIR 原则的适配，提供的 DMP 组件为数据概述、数据的 FAIR 原则、资源配置、数据安全、数据伦理及其他问题，其评估准则（*Horizon 2020 FAIR DMP Evaluation Rubric*）[46]面向科研教育机构、评审员等主体提供服务，评估的要素为 DMP 编制中是否覆盖全数据生命周期管理规划。

表 9　CR、Criterion 与 FAIR 原则的兼容性[20]

CR			Criterion		
处理阶段	指标编号	与 FAIR 原则的兼容性	处理阶段	指标编号	与 FAIR 原则的兼容性
一般信息	Null	/			
1 数据描述和现有数据的收集或再利用	1a	R1.2	1 提供持久且唯一的标识符（PIDs）	1a	F1，F3，A1
	1b	I1，R1.2，R1.3		1b	F1，F3，A1
2 文档与数据质量	2a	F2，I1，I2，I3，R1，R1.3		1c	F1，F3，A1，R1.2
	2b	I2，R1，R1.2	2 元数据	2a	F2，F3，F4，A1，A1.1，I2，R1，R1.2
3 科研过程中的存储与备份	3a	/		2b	F2，F3，F4，A1，A1.1，I2，I3，R1，R1.2
	3b	/		2c	F2，F3，F4，A1，A1.1，A2，I2，R1，R1.2
4 法律和道德要求、行为准则	4a	/		2d	F2，F3，F4，A1，A1.1，I1，I2，R1，R1.2，R1.3
	4b	A1.2，R1.1		2e	F2，F3，F4，A1，A1.1，I2，R1，R1.2
	4c	A2	3 数据访问和使用许可协议	3a	A1.2
5 数据共享与长期保存	5a	A1.2，A2，I3，R1.1		3b	A1.2，R1.2
	5b	/		3c	A1.2
	5c	A1，A1.1，A1.2，I3		3d	A1.2，R1.1
	5d	F1，F3，A2		3e	A1.2
6 数据管理职责与资源	6a	/	4 保存	4a	R1.2
	6b	/		4b	/

同样采用模本法归纳“整合 DLC-SH”和“机器可操作”两种导向的共同点，模本主客体选择标准如下：由表 9 可知，政策 SE[20] 体现了两种导向的良好兼容，为确保主客体模本一致，以其为主体模本，以其 CR、Criterion 指标与 FAIR 原则的兼容及其实践为客体模本进行分析。如表 10 所示，国际科研数据政策的两种导向在基础环境、实现途径和最终目标上基本一致。未来，人机协同将是科研数据管理与共享的基本趋势，全球协同、善治是终极目标。“整合 DLC-SH”和“机器可操作”导向下的科研数据管理与共享工作存在共同的支撑基础和社会分工，互相融合以提升效能成为主流趋势。

6.2　国际科研数据政策导向启示

6.2.1　强化政策的导向功用，融入国际工作体系

欧盟地平线 2020 计划资助的 FOSTER Plus 项目指出：开放存取与开放数据是开放科学运动的两支主要力量[47]。在开放存取、开放科学运动持续深入发展的时代背景下，开展良好的科研数据管理与共享国际协作，既是发展“数字新基建”的必然要求，更是开放科学交流与合作的现实需要。国际上，科研数据政策标准的效能已深入到微观层面，如澳大利亚[1]、英国[2]和瑞士[3]发布的相关政策影响着欧美国家和地区的科研数据管理与共享实践，助力其形成面向多元利益相关者、覆盖全数据生命周期的政策保障体系；欧盟《2020 计划框架下的科学出版物与科研数据开放获取管理指南》（政策 EC_1）、《2020 计划框架下的 FAIR 数据管理指南》（政策 EC_2）勾勒出科研数据管理及开放获取的框架，并提供了详细的指导意见[39]。我国《办法》第九条第二款虽规定：“科研院所、高等院校等法人单位具有建立健全本单位科学数据相关管理制度的职责”[4]，但仍缺乏与国际先进政策标准对接的机制，相关实施细则的实操性亦较为薄弱。从“整合 DLC-SH”和“机器可操作”导向下的国际实践可以看出，开展科研数据管理与共享的国际交流和合作尤为必要。

因此，应强化政策的导向功用，积极对接国际先进政策标准，探究分步实施细则，主动融入到国际科研数据管理与共享工作的合作体系当中。我国各利益相关者应积极探索科研数据管理与共享跨域、跨境的“双边、多边”合作模式，构建国际化数据合作机制与平台并制定相关规则；参考国际先进经验完善科研数据管

表 10　基于模本法的两种导向共同点

模本法流程	双证方向	两种导向的共同特征提取
提取主体模本特征	理论	政策 SE： ①基础环境－基础设施、法律政策、资金支持、工具平台和 DMP 模板等 ②实现途径－人机和谐共生 ③最终目标－构建全球善治的数据生态
提取客体模本特征	实践	CR、Criterion 与 FAIR 原则的兼容性： ①基础环境－数据基础设施，如“CoreTrustSeal＋FAIR”认证存储库 169 个 ②实现途径－人机协同，如 CR、Criterion 与 FAIR 原则的兼容项达 23 项，兼容率高达 74.19% ③最终目标－推动全球协同的科研数据管理与共享，实现“数据互联网”全球布局
主客体特征指向	①人机协同是基本趋势；②全球协同、善治是基本目标；③“整合 DLC-SH”和“机器可操作”导向存在互相融合的可能性	
结果	支持：两种导向存在相同之处	

理与共享政策的体系，并制定切实可行的实施细则。例如，在《办法》[4]以“数据生命周期为经、利益相关者为纬”的工作体制下，编制科研数据政策参考框架，制定能够有效适配 FAIR 原则、TRUST 原则等国际政策标准的实施细则，为 DMP 编制与评估，（元）数据创建与存储，数据存储（存档）与共享、出版与分发，数据分析与重用、评估与验证等工作提供细致的标准和操作指南。现阶段，我国部分省市和科研机构已制定了《办法》的实施细则，全国信息技术标准化委员会大数据标准工作组已颁行了 12 项大数据国家标准，《信息技术科学数据引用》[48]（GB/T 35294－2017）提供了“通用科学数据引用格式”和“基于 OID 的科学数据引用方式”，但科研数据政策标准的制定与完善工作仍道阻且长。未来，要打通科研数据管理与共享的生命周期流程，为多元利益相关者提供全面的政策标准、指南和工具资源支持；在保障数据主权安全的基础上，秉持科研数据管理与共享工作发展的整体观进行政策、制度设计，通过主动引入、创建依次递进的指导性原则与标准来推动实践工作的有序发展，积极参与全球竞争、合作和对话。

6.2.2　重塑工作机制与流程，完善数据治理格局

“整合 DLC-SH”的国际科研数据政策实践高度重视组织管理与业务流程的核心作用，“机器可操作”关注到了人机协同的重要性。例如，政策 SE[20]的文本组织及其规范的科研数据管理与共享工作思路为：服务于科研（资助）机构、科研人员和评审员等利益相关者，主要涉及 DMP 撰写与评估、可信数据存储两个方面，实质上还为全数据生命周期流程的工作提供详细指导，并与 FAIR 原则形成了良好兼容。科研数据管理与共享工作的核心利益相关者在科研数据政策制定、数据生命周期流程管理与服务、技术工具研发与应用、人才队伍建设等业务工作中负有不同的权责[34]，应根据“整合 DLC-SH”和“机器可操作”两种导向的特征理清工作思路，打造“组织管理科学、业务流程有序”的数据治理新格局。

（1）重塑“组织管理－业务流程”框架下的科研数据管理与共享工作机制与流程。一方面，采纳《办法》形成的“分级管理，分工负责”工作机制[4]，调整与完善服务分工，使各级各类主体能够各司其职，助力实现科研数据的高效管理与利用。可依据各类利益相关者的权责在实施细则、工作机制设计上进行规范和管理。例如，政府、科研（资助）机构、高校及其科研管理部门、图书馆等负责科研数据政策的制定与完善；图书馆、数据馆员等负责数据生命周期流程管理与服务；高校、图书馆、数据中心技术部门等负责技术平台和工具资源的研发与应用；高校、图书馆、数据中心等法人单位负责人才队伍建设；平台工作人员等负责可信存储库与平台的维护和运营；DMP 评审员、同行评议人员等负责数据技术与科学质量的评估和控制；科研人员等则负责数据的创建、共享与重用[34]，可参考国际准则以及《创建良好 DMP 的十个简单规则》[49]、《数字数据存储的十个简单规则》[50]等进行 DMP 编制、数据创建与存储。另一方面，重塑科研数据管理与共享的全生命周期服务流程。依据“整合 DLC-SH”导向优化计划、管理、共享和利用 4 个阶段关涉的 DMP 以及数据创建、存储、出版、共享、分析、重用、评估与验证等流程工作，我国应重点关注在 DMP 撰写与评审、数据出版、数据分析与重用、数据评估与验证等方面的薄弱环节，例如：将 DMP 纳入制度设计，并基于政策 SE、EC_2 等提供的或自主研发的 DMP 评估准则明确核心评估领域、划分评估等级，构建成熟度评估模型和评价指标体系，以便进行更科学的 DMP 评估，引领全数据生命周期工作不断革新。

（2）完善国家数据治理格局。要求搭建科研数据管理与共享工作整体规划、统筹、协调、监督和保障的组织架构和制度安排，形成“协同合作，横纵贯通”的

工作机制，为数据治理平衡提供坚实的体制机制保障；营造良好的科研数据管理与共享工作发展基础环境，如在数据基础设施建设、政策标准制定和服务引导上持续发力，提供学科数据存储库、通用存储平台、云备份平台和自主研发平台等可信存储库支持，并加强基础环境、DMP 以及数据收集、描述、命名、转移、安全、发布、清洗、挖掘、分析、可视化、引用与验证等方面的工具资源和标准规范的研发、供给和应用。

6.2.3 “人文与技术”相结合，构建“人机协同”生态

“人机协同”是“整合 DLC-SH”和“机器可操作”导向的主要共同点。譬如，政策 SE[20] 的实现高度依赖于 DMP OPIDoR、DMPTuuli 和 DMPonline 等 DMP 工具以及科研（资助）机构提供或科研人员自主识别的可信存储库和科研数据管理平台的支持；为有效实现 FAIR 原则，可采用数据共享试验阶段联盟开发的数字资源评估工具“FAIRshake”及其指标、标准来评测数据资源的可发现性、可访问性、互操作性和可重用性[51]。由此可见，应以“人文与技术”相结合，实现围绕人类所建构的“人际秩序”和围绕智能技术所建构的“机器秩序”有机融合，构建“人机协同”的科研数据管理与共享生态。重点任务如下：

（1）引入科研数据管理与共享的国际准则。我国在制定相关政策、标准，开发相应软件和工具资源时，要适度吸纳国际科研数据政策、指导性原则的精神。例如，吸收 FAIR 原则的先进理念，采取“自下而上的原则采纳、协调统一的合作推进、在已有成果和经验基础上推动原则实施和由内而外推动国际合作”[13] 等方式推进 FAIR 原则，确保全国范围内的科研数据管理与共享工作实现机器友好以及在技术、法律等层面的互操作，保障多元利益相关者在科研数据管理、共享与使用等方面的权利；采纳“FAIRsFAIR 数据对象评估 + FAIR 数据成熟度模型”方案设计自主的、自动化的和可操作的评估模型与工具，为数据建设和数据管理过程中利益相关者评估 FAIR 数据的遵循度提供依据[52]；充分考虑政策 SE 对 DMP 核心要求、可信存储库选择标准等方面的规范以及其与 FAIR 原则的兼容性，推动与之相适配的标准规范、工具平台的研发和应用。

（2）健全数据基础设施，融入全球“数据互联网”。现阶段，我国在数据基础设施建设上还相对落后，如在 Re3Data 整合的全球 2 416 个数据存储库中，中国仅有 41 个（排名全球第 11 位）[53]；CoreTrustSeal + FAIR 认证下的可信存储库仅有 12 个（含港澳台）[43]；在 RDA 编制的全球元数据标准目录中亦鲜有贡献。我国应乘势而为，与时代和世界发展同频共振，强化数据标准化工作，加快国际化进程，主动融入全球“数据互联网”建设中去；通过完善数据基础设施，着力建成融“数据、技术、人员和规则”于一体的科研数据管理与共享生态体系[54]。

7 结语

质量可靠的科研数据是科研过程的关键组成部分，且应在安全可控的前提下永久、公开和免费提供，以便重复使用，发挥其最大化的经济、社会和科学价值。为充分挖掘数据价值和实现科研数据管理与共享工作的有序运行，国际上形成了丰富的科研数据政策和实践，本文将其导向归纳为“整合 DLC-SH”和“机器可操作”两种模式。通过对两种政策导向的特征进行系统梳理、分析和思考，发现两者之间虽有不同，但是存在相互融合的主要趋势。同时，科研数据政策导向作为一种重要原则，主要功用在于指引实践工作的发展。在后续研究中，应当进一步解构两种政策导向与我国科研数据管理与共享工作政策设计、管理体制机制的适配性，探究将其与信息技术、政策法规等其他手段有机结合的数据治理之道。

参考文献：

[1] ANDS. Outline of a research data management policy for Australian universities/institutions[EB/OL]. [2022 - 01 - 14]. https://rdc-drc. ca/wp-content/uploads/Institutional-Research-Data-Management-Policies. pdf.

[2] DCC. Five steps to developing a research data policy[EB/OL]. [2022 - 01 - 14]. http://www. dcc. ac. uk/sites/default/files/documents/publications/DCC-FiveStepsToDevelopingAnRDMpolicy. pdf.

[3] DLCM. Research data management policy template[EB/OL]. [2022 - 01 - 14]. https://www. dlcm. ch/download_file/force/68/372.

[4] 国务院办公厅. 关于印发科学数据管理办法的通知[EB/OL]. [2022 - 01 - 14]. http://www. gov. cn/zhengce/content/2018-04/02/content_5279272. htm.

[5] 邢文明，洪芳林，李晓妍. 科学数据管理体系的二维视角——《科学数据管理办法》解读[J]. 图书情报工作，2019，63(23)：30 - 37.

[6] GO FAIR. FAIR principles[EB/OL]. [2022 - 01 - 14]. https://www. go-fair. org/fair-principles/.

[7] 邢文明，杨玲. 中美科学数据政策比较——以《科学数据管理办法》和《促进联邦资助科研成果获取的备忘录》为例[J/OL]. 图书馆论坛[2022 - 01 - 14]. http://kns. cnki. net/kcms/detail/44. 1306. G2. 20210621. 1728. 007. html.

[8] HODSON S, MOLLOY L. Current best practice for research data management policies[R]. Paris:CODATA,2020:1－19.

[9] CORREA DA SILVA F C. Infrastructure and international development policies for data management research[J]. Biblios-revista de bibliotecologia y ciencias de la informacion,2016,63:44－55.

[10] 中国科协创新战略研究院. 欧洲实施科研数据管理政策:“科学欧洲”成员机构经验做法[R]. 北京:创新研究报告,2020.

[11] 王丹丹,董金金,杨嘉敏.《科研数据管理国际联盟实用指南》研究及启示[J]. 数字图书馆论坛,2021(4):17－24.

[12] 宋佳,温亮明,李洋. 科学数据共享 FAIR 原则:背景、内容及实践[J]. 情报资料工作,2021,42(1):57－68.

[13] 翟军,梁佳佳,吕梦雪,等. 欧盟开放科学数据的 FAIR 原则及启示[J]. 图书与情报,2020(6):103－111.

[14] 邱春艳. 开放科学愿景下欧盟推进 FAIR 原则的路径、经验及启示[J]. 情报理论与实践,2021,44(5):199－205.

[15] 孟祥保,张璇. OCLC《科研数据管理的现实》系列报告解读与启示[J]. 图书情报工作,2019,63(7):38－46.

[16] 尤霞光,盛小平. 8 个国际组织科学数据开放共享政策的比较与特征分析[J]. 情报理论与实践,2017,40(12):40－45.

[17] 盛小平,王毅. 利益相关者在科学数据开放共享中的责任与作用——基于国际组织科学数据开放共享政策的分析[J]. 图书情报工作,2019,63(17):31－39.

[18] 唐义,张晓蒙,郑燃. 国际科学数据共享政策法规体系:Linked Science 制度基础[J]. 图书情报知识,2013(3):67－73.

[19] 邢文明. 我国科研数据管理与共享政策保障研究[D]. 武汉:武汉大学,2014:1－198.

[20] Science Europe. Practical guide to the international alignment of research data management-extended edition[R]. Brussels: Science Europe AISBL,2021:1－53.

[21] Research Data Alliance. The TRUST principles:an RDA community effort[EB/OL]. [2022－01－14]. https://www.rd-alliance.org/trust-principles-rda-community-effort.

[22] CODATA. The Beijing declaration on research data[EB/OL]. [2022－01－14]. http://www.codata.org/uploads/Beijing%20Declaration-19-11-07-FINAL.pdf.

[23] Research Data Alliance. CARE principles for indigenous data governance[EB/OL]. [2022－01－14]. https://www.gida-global.org/care.

[24] European Research Council. Guidelines on implementation of open access to scientific publications and research data[EB/OL]. [2022－01－14]. https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf.

[25] OCLC. The realities of research data management[EB/OL]. [2022－01－14]. https://www.oclc.org/research/publications/2017/oclcresearch-research-data-management.html.

[26] European Commission. Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020[EB/OL]. [2022－01－14]. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.

[27] European Commission. Guidelines on FAIR data management in Horizon 2020[EB/OL]. [2022－01－14]. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.

[28] International Science Council. Open data in a big data world[EB/OL]. [2022－01－14]. https://council.science/publications/open-data-in-a-big-data-world/.

[29] World Data System. Data sharing principles[EB/OL]. [2022－01－14]. https://www.worlddatasystem.org/files/WDS_Data_Sharing_Principles_2015.pdf.

[30] 叶鹰. 试论图书情报学的主干知识及有效方法:兼论双证法和模本法之效用[J]. 中国图书馆学报,2021,47(3):58－66.

[31] 金贞燕,阿童木. 科研数据管理服务内容体系构建研究[J]. 情报理论与实践,2021,44(8):42－50.

[32] FREEMAN R E. Strategic management: a stakeholder approach[M]. Marshfield:Pitman,1984:31－32.

[33] STEELEWORTHY M. Research data management and the canadian academic library: an organizational consideration of data management and data stewardship[J]. The Canadian journal of library and information practice and research,2014,9(1):1－11.

[34] 秦顺. 面向一流高校建设的图书馆科研数据管理服务研究——以整合 DLC-SH 为视角[J]. 图书情报工作,2021,65(4):28－39.

[35] ALAN H. From Bethesda to Beijing - open research data has arrived! [EB/OL]. [2022－01－14]. https://figshare.com/blog/From_Bethesda_to_Beijing_-_Open_Research_Data_has_arrived_/536.

[36] CARROLL S R,HERCZOG E,HUDSON M,et al. Operationalizing the CARE and FAIR principles for indigenous data futures[J]. Scientific data,2021,8(1):1－6.

[37] International Science Council. Science international to agree international accord on open data[EB/OL]. [2022－01－14]. https://council.science/current/news/science-international-to-agree-international-accord-on-open-data/.

[38] FORCE11. Joint declaration of data citation principles[EB/OL]. [2022－01－14]. https://www.force11.org/datacitationprinciples.

[39] 张玉娥,王永珍. 欧盟科研数据管理与开放获取政策及其启示——以“欧盟地平线 2020”计划为例[J]. 图书情报工作,2017,61(13):70－76.

[40] 张文萍,宋秀芬,魏银珍,等. 基于 FAIR 标准的科学数据融合体系研究[J]. 中国图书馆学报,2020,46(6):41－54.

[41] WILKINSON M D,DUMONTIER M,AALBERSBERG I J J,et al. The FAIR guiding principles for scientific data management and stewardship[J]. Scientific data,2016,3(1):1－9.

[42] FAIRsFAIR. CoreTrustSeal + FAIR DPC FAIR forever webinar[EB/OL]. [2022－01－14]. https://www.dpconline.org/docs/miscellaneous/events/2021-events/2428-2021-03-18-dpc-fair-for-

ever-hlh-fairsfair-coretrustseal-v01-00/file.
[43] CoreTrustSeal. Core certified repositories[EB/OL].[2022-01-14]. https://www.coretrustseal.org/why-certification/certified-repositories/.
[44] FAIRsFAIR. Recommendations on practice to support FAIR data principles[EB/OL].[2021-01-14]. https://www.fairsfair.eu/recommendations-practice-support-fair-data-principles.
[45] FAIRsFAIR. FAIRsFAIR data object assessment metrics: request for comments[EB/OL].[2025-01-14]. https://www.fairsfair.eu/fairsfair-data-object-assessment-metrics-request-comments.
[46] EMILIE H, JAN L. Horizon 2020 FAIR DMP evaluation rubric [EB/OL].[2022-01-14]. https://bibliotheek.uhasselt.be/sites/default/files/uploads/RDM/H2020FAIRDMP_Rubric.pdf.
[47] FOSTER Plus. Open science[EB/OL].[2022-01-14]. https://www.fosteropenscience.eu/foster#taxonomy.
[48] 中华人民共和国国家质量监督检验检疫总局,中国国家标准化管理委员会.信息技术 科学数据引用[EB/OL].[2022-01-14]. http://c.gb688.cn/bzgk/gb/showGb? type = online&hcno = A495CA355BAF00D962AA8DD84C3B2C16.
[49] MICHENER W K. Ten simple rules for creating a good data management plan[J]. PLoS computational biology, 2015, 11(10): 1-9.
[50] HART E M, BARMBY P, LEBAUER D, et al. Ten simple rules for digital data storage[J]. PLoS computational biology, 2016, 12(10): e1005097.
[51] FAIRshake. JSON-LD enabled automated assessments[EB/OL].[2022-01-03]. https://fairshake.cloud/documentation/json-schema/.
[52] 叶兰. FAIR 数据评估模型与工具研究[J]. 图书情报工作, 2021, 65(16): 138-147.
[53] European Commission. Facts and figures for open research data [EB/OL].[2022-01-03]. https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/facts-and-figures-open-research-data_en#what-is-open-research-data.
[54] 李云婷,温亮明,张丽丽,等.科学数据共享系统的现状与趋势[J].农业大数据学报,2019,1(4):86-97.

作者贡献说明:
秦顺:提出研究思路,确定研究框架,撰写论文与定稿;
戴柏清:收集、整理和分析资料,参与论文修改。

Analysis of International Research Data Policy Guidelines

Qin Shun[1] Dai Baiqing[2]
[1] School of Information Management, Wuhan University, Wuhan 430072
[2] School of Information Resource Management, Renmin University of China, Beijing 100872

Abstract: [**Purpose/Significance**] Analyzing the orientation and characteristics of international research data policy can provide an important basis for improving research data policy guarantee system and optimizing research data management and sharing work in China. [**Method/Process**] This article summarized the two orientations of international research data policy which named as "Integrated DLC-SH" and "Machine Actionable" by using the methods of case analysis and literature research, and analyzed and verified the similarities and differences between the two orientations by using the method of model sample. [**Result/Conclusion**] It is found that the orientation of "Integrated DLC-SH" adheres to the overall view of research data management and sharing, follows the basic framework of "organization management - business process", and aims to build a global collaborative data governance general plan. The orientation of "Machine Actionable" provides some guiding principles of sequential progression, and attaches importance to the auxiliary role of the data infrastructure, which is in order to realize the global layout of "Internet of Data". The two orientations are different in theoretical basis, core components and service objects, but there are basically the same in basic environment, implementation approaches and final goals, and the integration of the two orientations has become the mainstream trend at this stage. The useful enlightenment to China is drawn up, including: strengthening the guiding function of policy and integrating into the international working system, reshaping the working mechanism and process and improving the data governance pattern, and "combining humanities with technology" to build a "man-machine cooperation" ecology.

Keywords: research data policy data life cycle stakeholder machine actionable FAIR principles